



Bayesian data-model synthesis for biological conservation and management in Antarctica

Heather J. Lynch¹, Mathew Schwaller²

Chris Che-Castaldo¹, Grant Humphries¹, Michael Schrimpf¹

¹Stony Brook University Ecology & Evolution

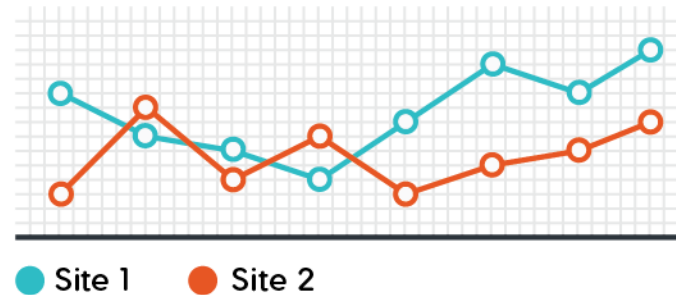
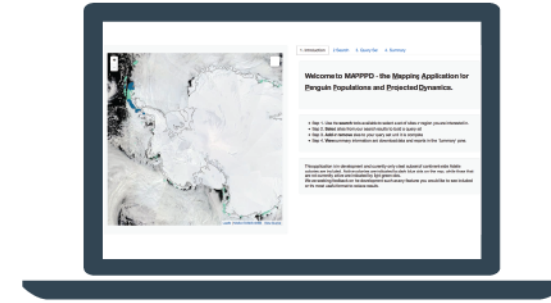
²NASA Goddard

Mapping Application for Penguin Populations and Projected Dynamics (MAPPPD)

Satellite imagery & Field counts

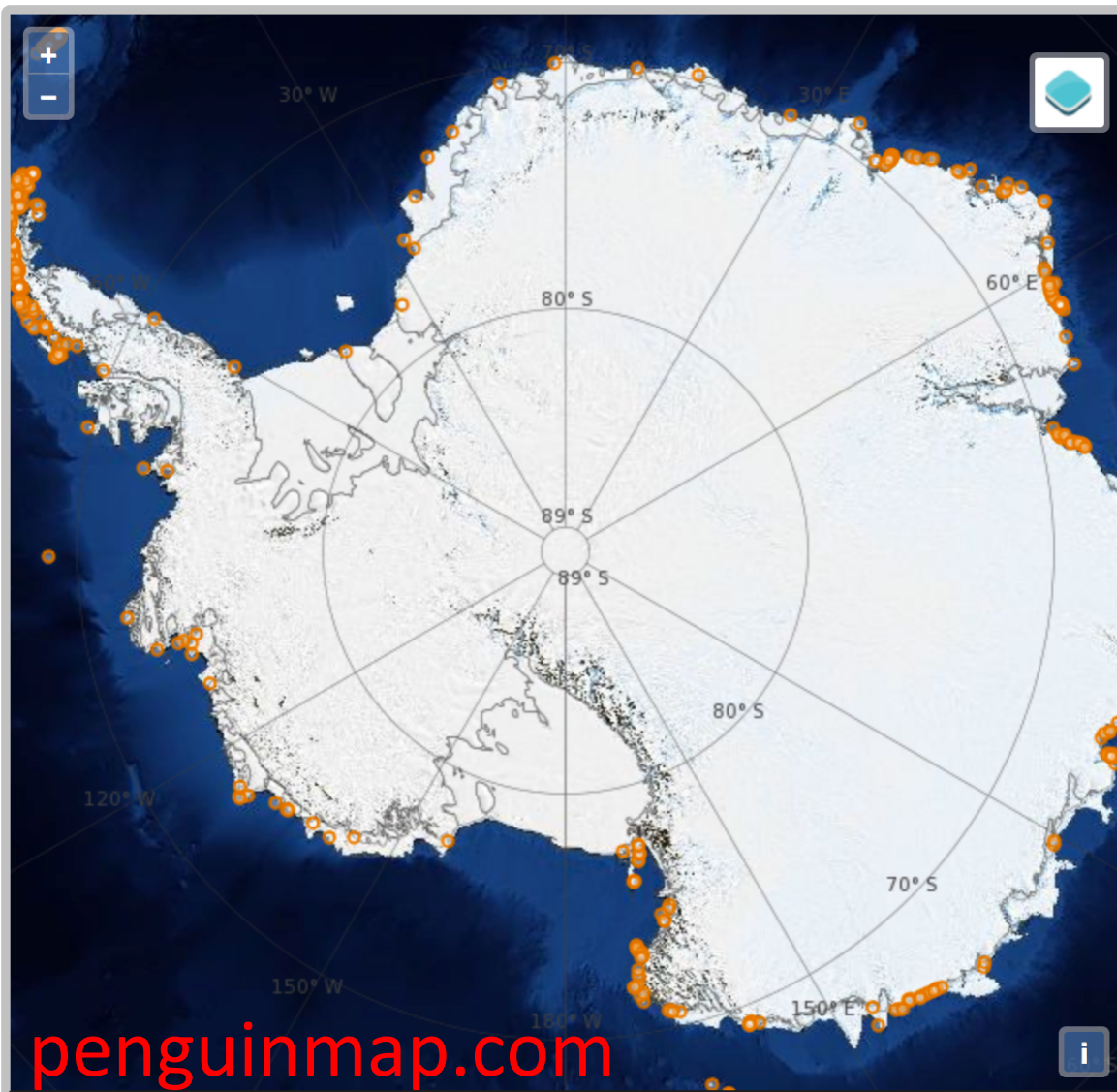


Map-based search engine for current abundance and predicted dynamics



Dynamic Naive Bayesian
Network Modelling

**“Data-to-policy”
pipeline**



Results

Welcome to MAPPPD, Version 1.2

Last updated: January 1, 2017

The MAPPPD project aims to deliver **open access** penguin population data for the Antarctic continent, and occupancy probabilities for flying birds around the Antarctic Peninsula.

Our database currently contains 3176 records of colony counts from 101 sources and 660 sites around the Antarctic.

MAPPPD contains counts for these species:

Adélie	1281
Chinstrap	846
Gentoo	932
Emperor	117

Find and download our metadata on the [Global Change Master Directory](#)

Explore the data by clicking on one of the search tools below the map to begin.

By accessing MAPPPD you agree to the [Terms of Use](#)

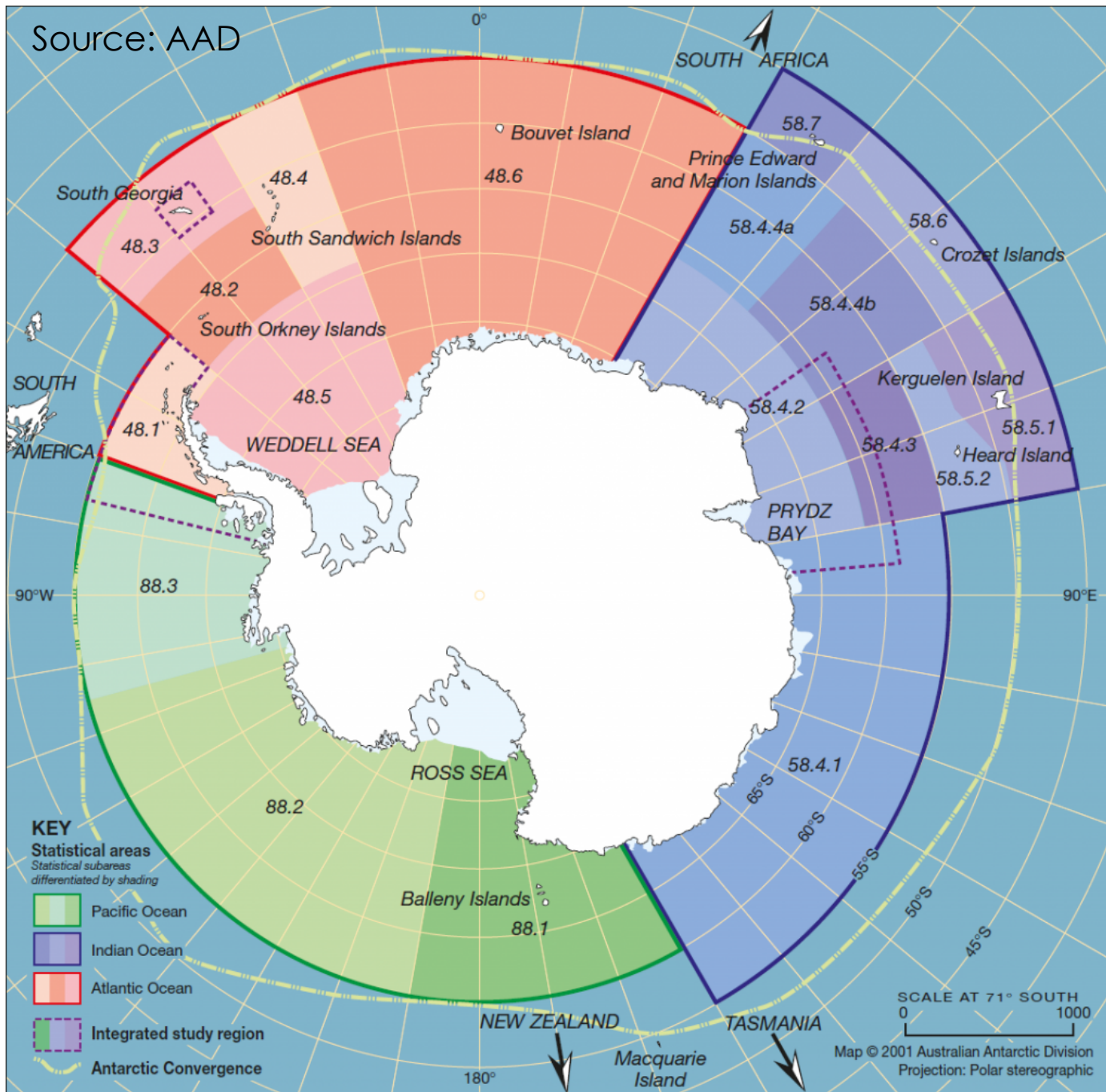
Upload Polygon

Search by Polygon

Search by Site

Search by CCAMLR region

Source: AAD



Now that we have MAPPPD,
what can we do?



Adélie

Study the efficacy of the Adélie
penguin as an indicator species for
adaptive management of
Antarctic fisheries.

What to do with patchy multi-site data?

SITE	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
1	X	X					X			
2			X	X	X					X
3		X								
4				X						
5		X						X	X	
6			X							
7								X		
8						X				X
9					X					
10	X		X	X	X	X	X	X		
11		X	X	X						X
12				X						
13									X	
14	X			X		X	X			
15				X				X	X	X
16						X				

Need models to estimate the unknown population in years of missing data ('latent states').

This is what we are usually doing...

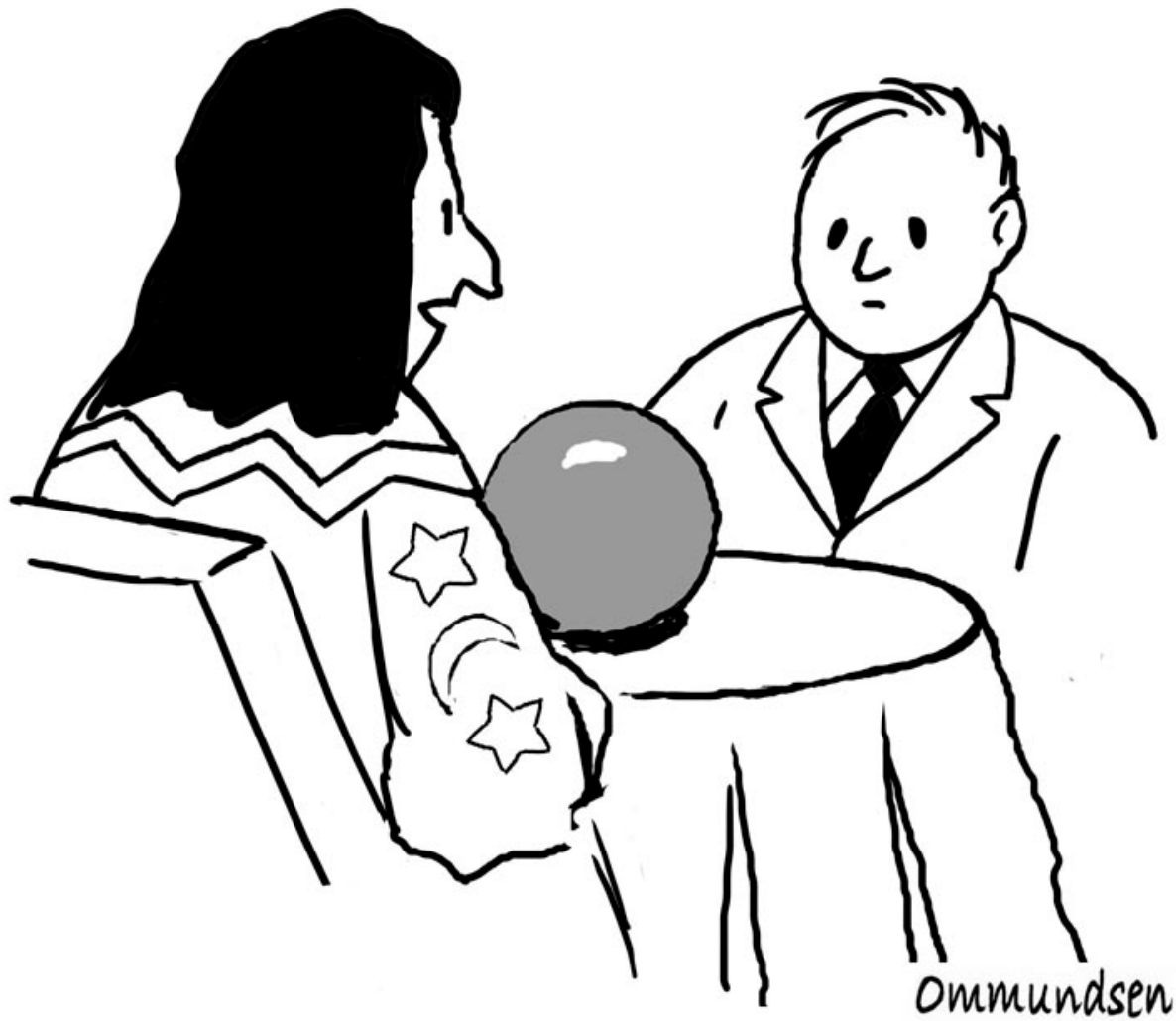
What to do with patchy multi-site data?

SITE	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
1	X	X					X			
2			X	X	X					X
3		X								
4				X						
5		X						X	X	
6			X							
7								X		
8						X				X
9					X					
10	X		X	X	X	X	X	X		
11		X	X	X						X
12				X						
13									X	
14	X			X		X	X			
15				X				X	X	X
16						X				

Need models to estimate the unknown population in years of missing data ('latent states').

This is what we are usually doing...

...but this is often what we need.



Adélie Abundance ~ Covariates

268 Adélie populations
around Antarctica

**“Is this needed for a
Bayesian analysis?”**

Process model

“true” abundance = $f(\text{biological covariates, process noise})$

Observation model

Observed abundance = $g(\text{“true” abundance, in situ precision})$

X	X	X	
			X
X	X		
			X
X		X	
		X	X
X			
	X		X
X		X	

Observed counts



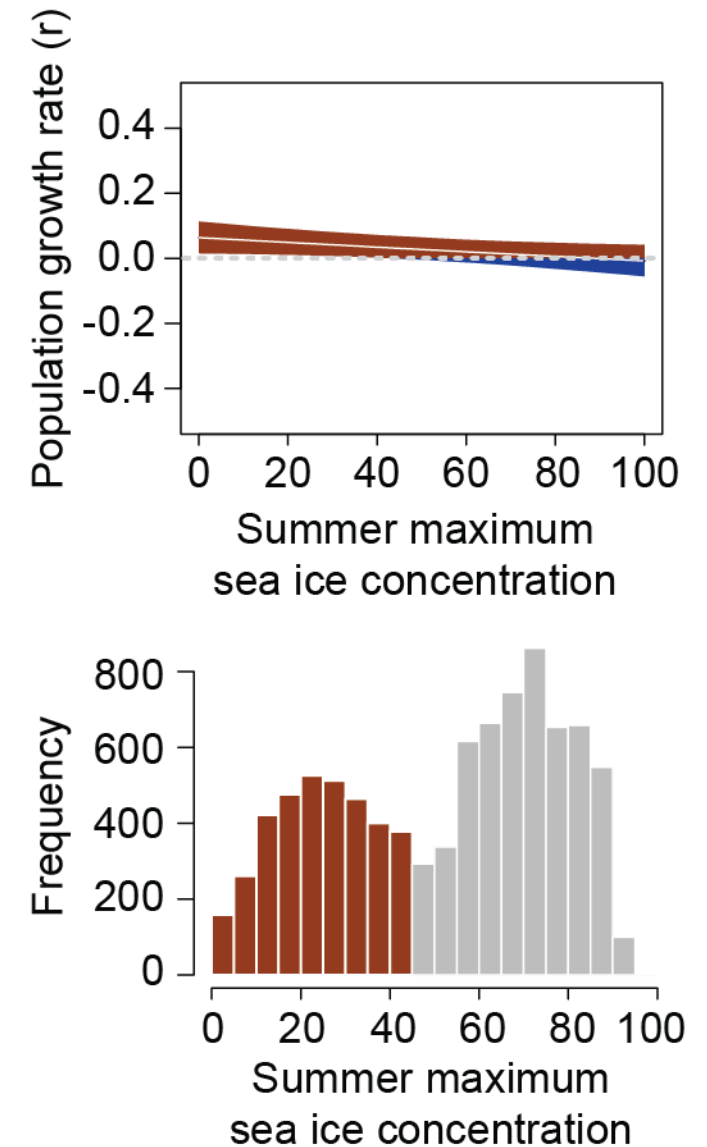
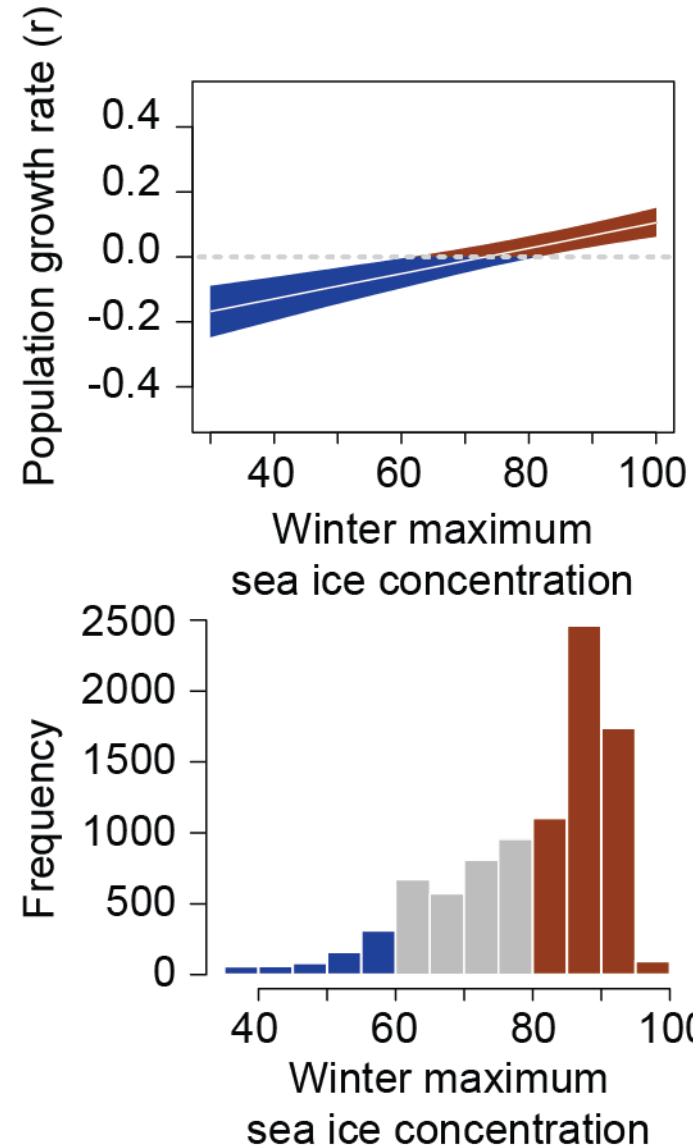
$$N_{t+1} = f(N_t)$$

Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z
Z	Z	Z	Z

“true” abundance

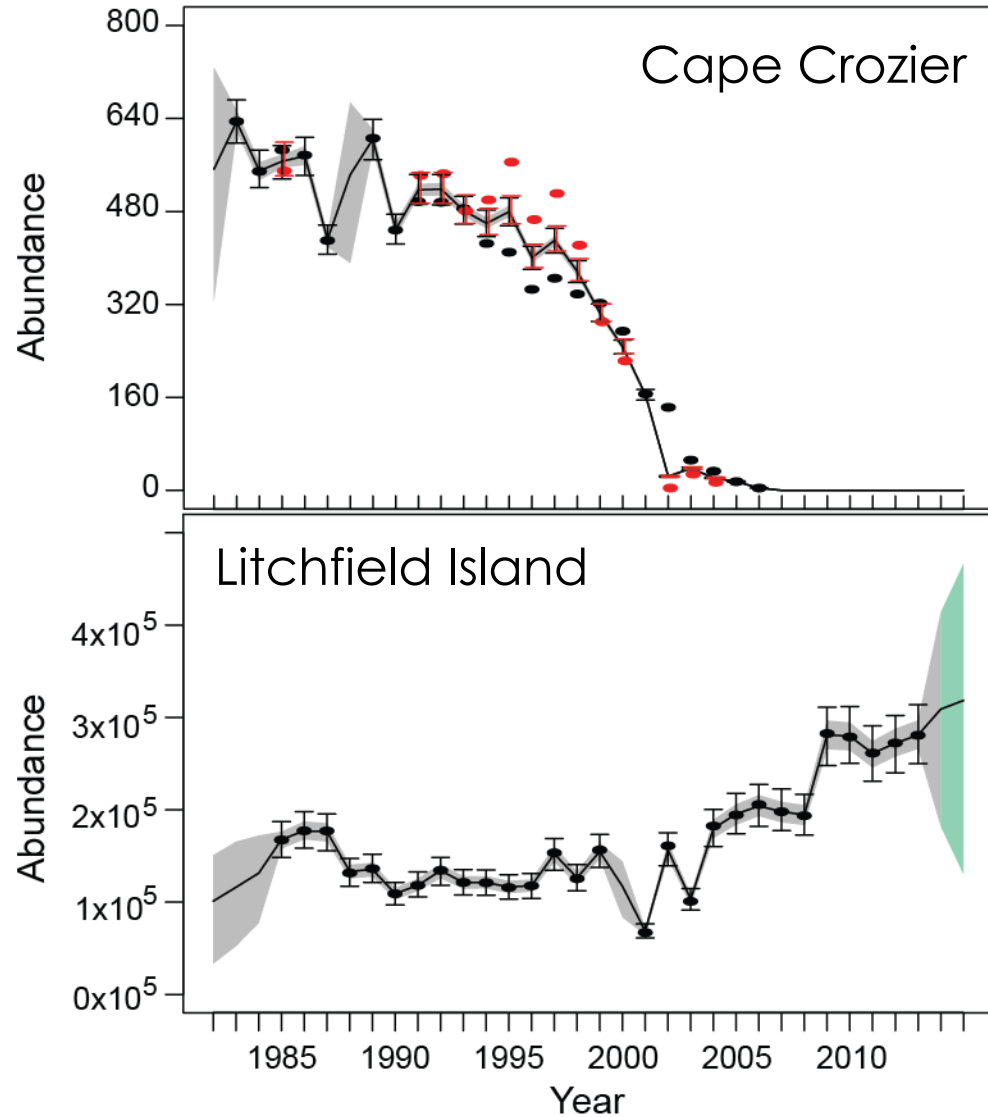
What did we find?

- 1) Interannual growth rates positively associated with maximum winter sea ice in the previous several years and negatively associated with maximum summer sea ice in $t-4$
- 2) **Almost all of the interannual variability in growth rates remains unexplained**

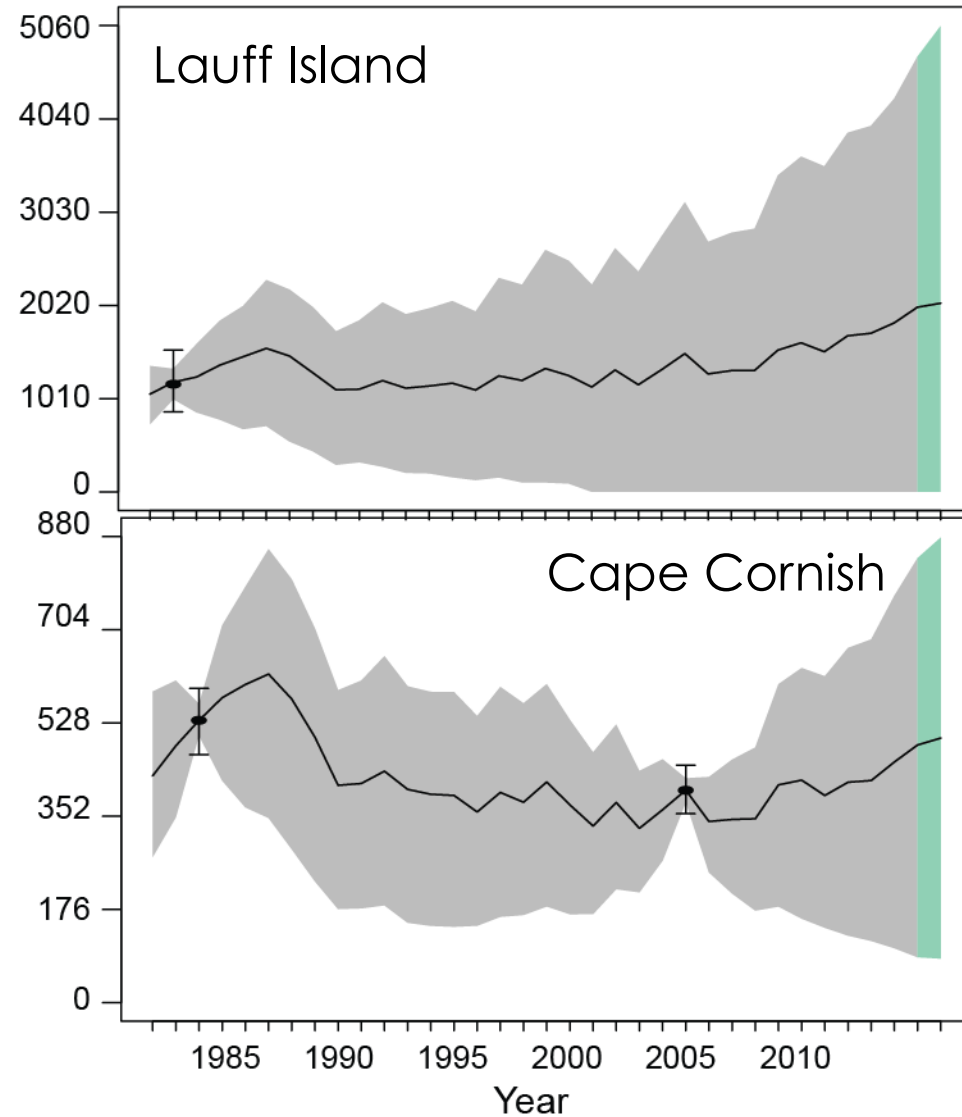


At the site level

Data-rich sites



Data-poor sites

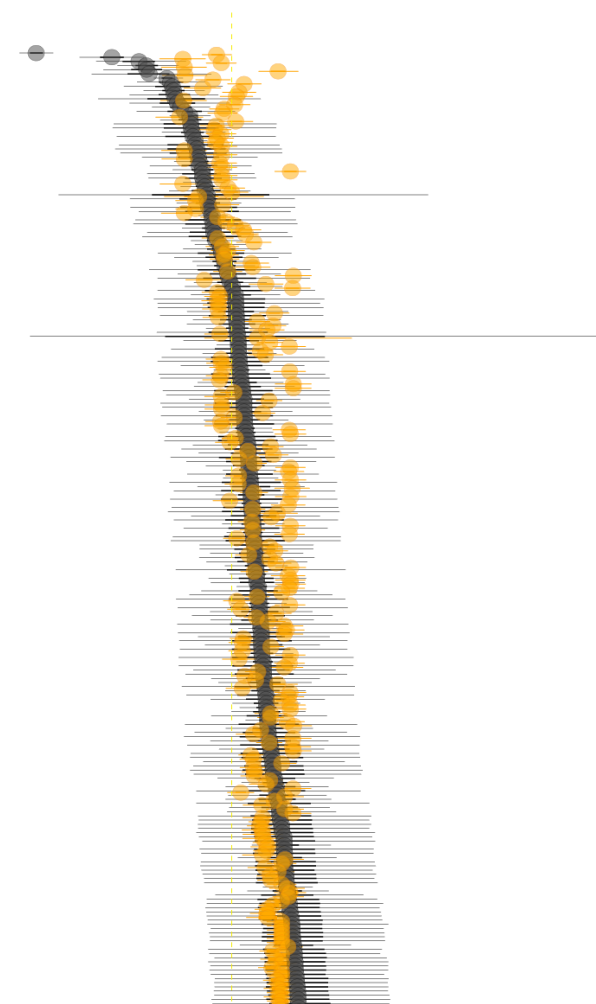
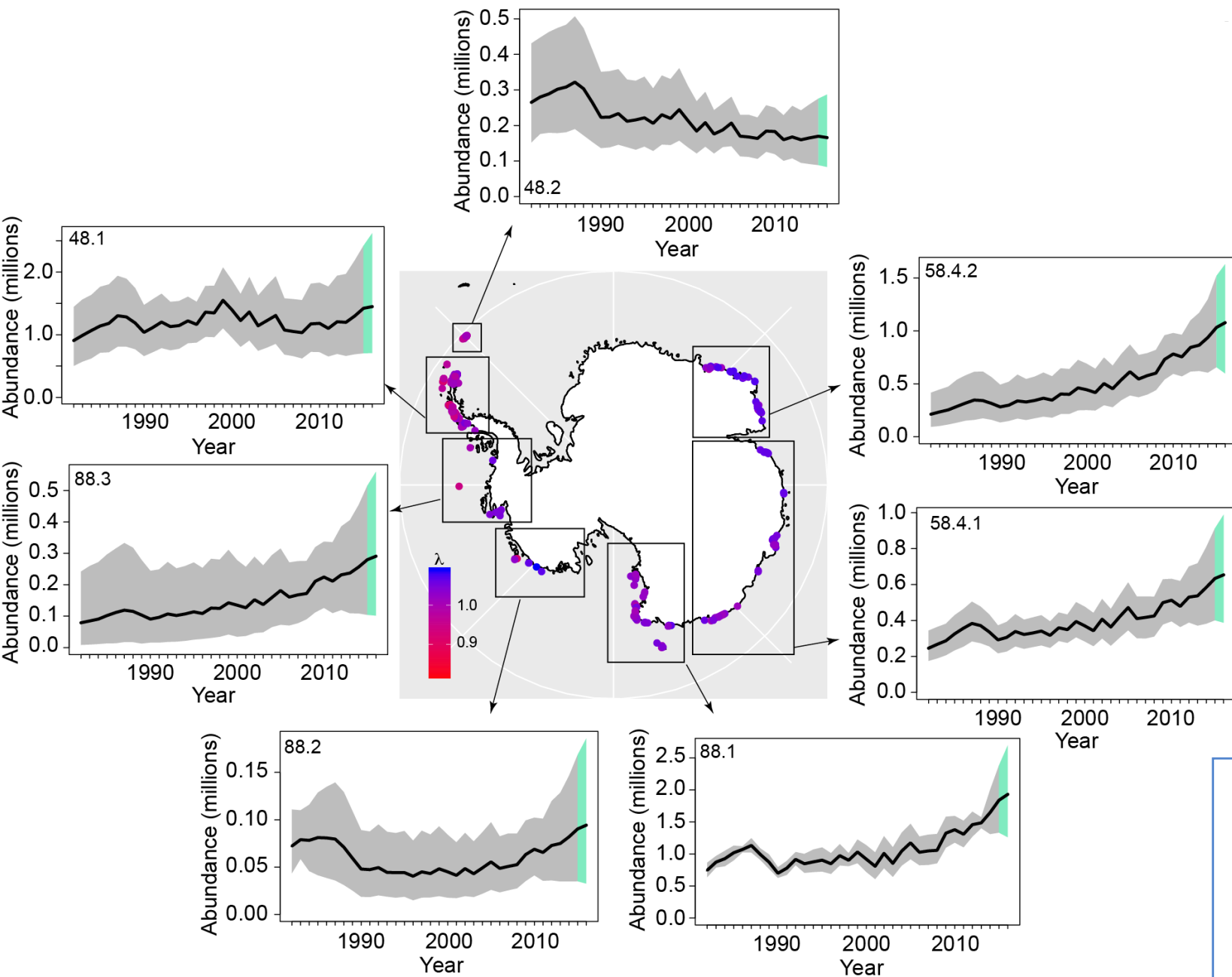


How does the model
infer abundance
when there is no
data?

Shared covariates
allow for a 'best-
guess' in years with
missing data.

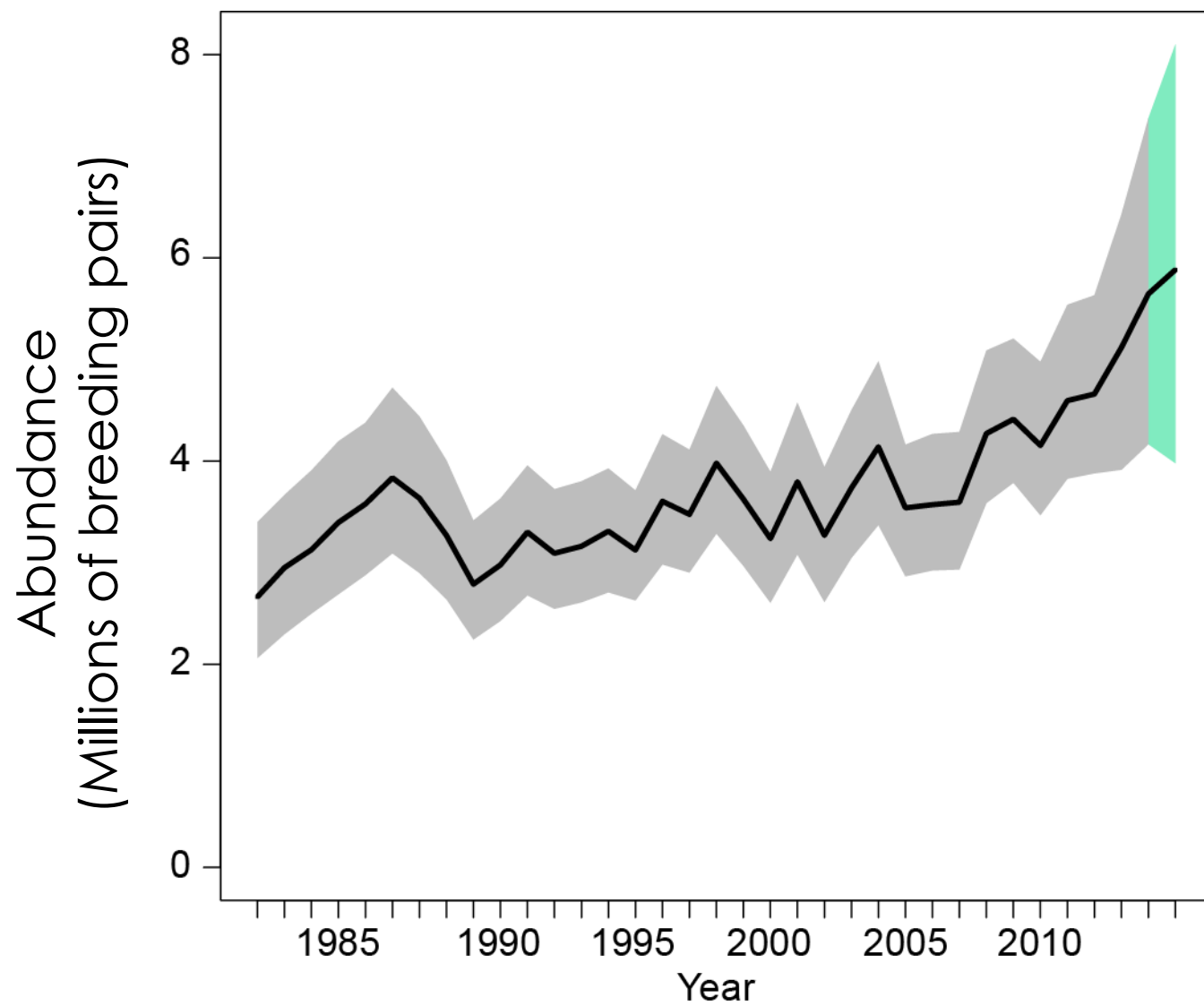
Est. from nest counts in black

Est. from chick counts in red



Remarkably consistent with Cimino et al. (2016) who inferred abundance trends from habitat suitability models fit using presence – absence data

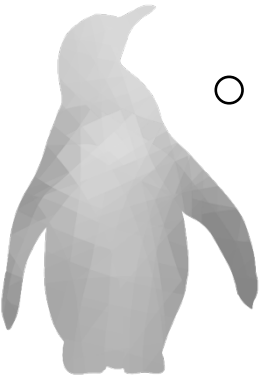
Everything south of 60° S



Where do we go from here?

- High-resolution remote sensing is answering questions we were not even asking 5-10 years ago (and not just for penguins!), and MAPPPD is a key component of sharing this new data stream with policymakers
- No longer have to trade spatial resolution for spatial extent
- We now have software that can provide Antarctic stakeholders quasi-real time information on abundance and distribution at any user-defined spatial scale (Population→SSMUs→Sub-areas→MPAs→Continent)
- Data rich model system for quantitative and spatial ecologists

(but we need to get them engaged with the data)



Random Walk of the Penguins

HOSTED BY DRIVEN DATA

[HOME](#)

[PROBLEM DESCRIPTION](#)

[ABOUT](#)























\$16,000.00

LEADERBOARD

You're not part of this competition. Yet...

[Join the competition!](#)

Penguins are among the most charismatic animals in the world and have captured the imaginations of news-makers, scientists, film producers, and the general public. Beyond their general intrinsic value, they are considered important ecosystem indicators. In other words, monitoring these beautiful species can tell us a lot about the general health of the Antarctic because penguins are important krill and fish predators, and changes (natural or anthropogenic) that influence prey abundance and environmental conditions will ultimately be detected through changes in distribution or population size.

	User or team	Best Score ⓘ ▲	Timestamp ⓘ	▲ Trend	# Entries ⇅
	jackh	4.8195	May 5, 2017, 1:43 a.m.		8
	mlandry	5.0319	May 8, 2017, 12:35 a.m.		6
	mathurin	5.0407	May 10, 2017, 5:56 a.m.		12
	leonardo	5.2104	May 17, 2017, 12:26 p.m.		2
	ambarishg	5.2132	May 18, 2017, 2:40 a.m.		4
	alexp910	5.2136	April 30, 2017, 2:14 a.m.		5
	QuantumDamage	5.2136	May 15, 2017, 12:13 p.m.		6
	bicarrio	5.2136	May 17, 2017, 1:44 a.m.		2
	madjuice	5.2136	May 18, 2017, 11:16 a.m.		2
	priya.ana	5.2273	May 16, 2017, 5:39 a.m.		6

A data science competition is a flavor of citizen science that we had not initially anticipated, but an amazing way to spread the word and get some new models for our decision support tool.

In just the first three weeks, **414 people** have signed up to participate and **208 unique models** already submitted.

What we've completed to date:

- Underlying PostgreSQL database for all four Antarctic penguins
- Front end application essentially complete
- Underlying population dynamics model for Adélie penguins is complete
- Archival Landsat survey (L4,L5,L7,L8) complete
- Characterization of ground targets with field spectrometer complete
- Occupancy model for non-penguin Antarctic seabirds complete and displayed in MAPPPD
- MAPPPD is currently in use for the community

Thanks NASA Ecosystem Forecasting
(and extra special thanks for supporting
the data science competition!).

Questions??

What we will do in the next year:

- Add ensemble model forecasts
- Integrate Landsat retrievals and clean-up workflows for transition to partner organization

